Università degli Studi del Molise

Dipartimento di Scienze Economiche, Gestionali e Sociali (SEGeS)

ECONOMICS & STATISTICS DISCUSSION PAPER

No. 065/12

A copula-based analysis of false discovery rate control under dependence assumptions

Roy Cerqueti

Mauro Costantini

Claudio Lupi

A copula-based analysis of false discovery rate control under dependence assumptions

Roy Cerqueti, Mauro Costantini, Claudio Lupi

Abstract

The false discovery rate (FDR) first introduced in Benjamini and Hochberg (1995) is a powerful approach to multiple testing. Benjamini and Yekutieli (2001) proved that the original procedure developed for independent test statistics controls the FDR also for positively dependent test statistics. Furthermore, Yekutieli (2008) showed that a modification of the original procedure can be used even in the presence of non-positively regression dependent test statistics. In this paper we elaborate on Yekutieli (2008) and introduce suitable classes of copulas to identify the conditions under which the dependence properties needed to control the FDR are satisfied.

Keywords: Multiple testing, False discovery rate, Copulas.
JEL: C12, C40.

1. Introduction

When many hypotheses are tested simultaneously, the risk of falsely rejecting truly null hypotheses increases dramatically. In one single test we usually reject the null if the test p-value, p, is such that $p < \alpha$, for a pre-specified level α. Since $p \sim \mathrm{U}_{(0,1)}$ under the null, we have that $\Pr(p < \alpha | \mathrm{H}_0) = \alpha$. But when $m \gg 1$ hypotheses are tested simultaneously it is likely that *at least* one of the p-values is less than α even if all the hypotheses are truly null. On the other hand, a researcher would probably like to identify as many "discoveries" as possible (Sorić 1989), while incurring in a small proportion of false positives. This is the motivation of the concept of *False Discovery Rate* (FDR) introduced by Benjamini and Hochberg (1995). In plain words, the FDR is the expected value of the proportion of errors among the rejected hypotheses.

Benjamini and Yekutieli (2001) proposed a procedure to control the FDR at level q for all joint test statistics, under a positive dependence. The proposed strategy consists in the application of the Benjamini-Hochberg (BH) (Benjamini and Hochberg 1995) scheme at level $q \big/ \left(\sum_{i=1}^{m} i^{-1}\right)$. An even more general procedure, the separate subsets BH (ssBH) procedure, has been introduced recently in Yekutieli (2008) in order to deal with more general forms of dependence.

In order to fix the ideas, we offer here a brief explanation of how the ssBH procedure works.

Denote by $\boldsymbol{p} = (p_1, \ldots, p_m)'$ the vector of the m p values associated with the components of the collection of m statistics $\boldsymbol{t} = (t_1, \ldots, t_m)'$. Consistently with Yekutieli (2008), we assume that the p values in $\boldsymbol{p}$ are co-monotone transformations of the corresponding test statistics in $\boldsymbol{t}$. Divide $\boldsymbol{p}$ in S sub-vectors $\boldsymbol{p}^s$, for $s = 1, \ldots, S$. With a very intuitive notation, the statistics corresponding to $\boldsymbol{p}^s$ constitute a vector, that will be indicated with $\boldsymbol{t}^s$. Assume that the cardinality of $\boldsymbol{p}^s$ is m^s and denote as $\boldsymbol{p}_0^s$ the p values in $\boldsymbol{p}^s$ corresponding to the true null hypotheses. The level q ssBH procedure runs into two steps as follows:

1. For $s = 1, \ldots, S$, apply the BH procedure at level qm^s/m to test $\boldsymbol{p}^s$, and denote as $\boldsymbol{r}^s_{BH}$ the p values corresponding to the rejected hypotheses.

2. Reject the null hypothesis corresponding to $\boldsymbol{r}_{ssBH} = \bigcup_{s=1}^{S} \boldsymbol{r}^s_{BH}$.

This paper aims at developing formal arguments to characterise the relationship between FDR control and the dependence properties of the individual statistics involved, when the ssBH procedure advocated in Yekutieli (2008) is used. To achieve our aim, we model the stochastic dependence among the univariate statistics through the introduction of suitable families of copulas. In doing so, we are able to deal with dependence concepts more general than Pearson's correlation or those based on linearity, which are classically related to the limited world of normal random variables.

The rest of the paper is organized as follows. Section 2 concisely provides the statistical background. The main results are contained in Section 3. A discussion of the meaning and worthiness of the results is offered in Section 4. The proofs are collected in Section 5.

2. Statistical preliminaries

We stress that FDR control is strongly related to the stochastic dependence among the individual test statistics belonging to $\boldsymbol{t}$. Therefore, a detailed discussion on the dependence structure underlying the t's is needed. In this respect and to be self-contained, we recall here the concept of *positive regression dependency on each one from a subset* $I_0 \subseteq \{1, \ldots, m\}$ or, briefly, *PRDS on* I_0:

Definition 2.1. *Consider an increasing set*[1] D*. The vector* $\boldsymbol{t}$ *is assumed to satisfy the PRDS on* I_0 *if, for each* $i \in I_0$*, the conditional probability* $\Pr(\boldsymbol{t} \in D \,|\, t_i = x)$ *is nondecreasing in* x*.*

Benjamini and Yekutieli (2001) proved that the PRDS property on subsets of the test statistics t's corresponding to the true null hypothesis ensures the control of the FDR at a certain level by the BH procedure. Unfortunately, this result meets severe drawbacks in practice, because of the difficulty in showing the PRDS property. To overcome this problem, the *multivariate total positivity of order 2* or, briefly, *MTP2* property — a stronger dependence structure — can be used instead. We recall here the definition of MTP2:

Definition 2.2. *Let* f *be the joint density function of the* m*-variate random variable* $\boldsymbol{t}$*.* $\boldsymbol{t}$ *is said to be MTP2 if and only if, for each* $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *in* $\mathbb{R}^m$*, it results:*

$$f(\boldsymbol{x}) \cdot f(\boldsymbol{y}) \geqslant f(\min\{\boldsymbol{x}, \boldsymbol{y}\}) \cdot f(\max\{\boldsymbol{x}, \boldsymbol{y}\})$$

where the min *and* max *operators have to be intended* componentwise*.*

[1] A set D is said to be increasing when, if $x \in D$ and $y \geqslant x$, then $y \in D$.

A well known statistical result ensures that MTP2 $\implies$ PRDS on I_0 $\forall I_0$. Therefore, the dependence described by the MTP2 can be used instead of the PRDS on I_0, having in mind that the former condition is stronger.

Furthermore, the (linear) dependence between the individual t's in $\boldsymbol{t}$ can also be well represented by a non-diagonal variance-covariance matrix $(\sigma_{i,j})_{i,j=1,\dots,m}$, where variances are indicated with a unique index as: $\sigma_{i,i} = \sigma_i^2$. Hence, it is natural to guess a relationship between the value of the covariances and the MTP2 dependence property. In this respect, it is useful to recall a further definition of dependence between random variables.

Definition 2.3. *The random variables* $\{t_1, \dots, t_m\}$ *are associated if*

$$\mathrm{Cov}(g(t_1, \dots, t_m), h(t_1, \dots, t_m)) \geqslant 0,$$

for any coordinatewise nondecreasing functions $g, h : \mathbb{R}^m \to \mathbb{R}$ *for which this covariance exists.*

A classical result states that random variables that are MTP2 are also associated. Therefore, by using Definition 2.3 with $g(t_1, \dots, t_m) = t_i$ and $h(t_1, \dots, t_m) = t_j$, the relationship between the MTP2 property and covariance states immediately as follows:

Proposition 2.4. *Assume that* $\{t_1, \dots, t_m\}$ *are MTP2. Then* $\sigma_{i,j} \geqslant 0$*, for each* $i, j = 1, \dots, m$.

Note that Proposition 2.4 implies that if there exists a couple such that $\sigma_{i,j} < 0$, then $\{t_1, \dots, t_m\}$ are not MTP2.

A rather general way to capture the stochastic dependence structure among random variables is the introduction of the concept of *multivariate copula* (or, simply, *copula*). We now provide the definition of this concept, referring the reader to Nelsen (2006) for a detailed discussion.

Definition 2.5. *The function* $C : [0,1]^n \to [0,1]$ *is a copula if and only if:*
(C2.5.i) $C(u_1, \dots, u_n) = 0$ *if* $u_1 \times \dots \times u_n = 0$*;*
(C2.5.ii) $C(u_1, \dots, u_n) = u_{\bar{k}}$ *if* $u_k = 1$*, for each* $k \neq \bar{k}$*;*
(C2.5.iii) Given the n-dimensional rectangle $[a_1, b_1] \times \dots \times [a_n, b_n] \subseteq [0,1]^n$*, then*

$$\sum_{i_1=1}^{2} \cdots \sum_{i_n=1}^{2} (-1)^{i_1+\dots+i_n} C(u_{1,i_1}, \dots, u_{n,i_n}) \geqslant 0,$$

where $u_{j,1} = a_j$ *and* $u_{j,2} = b_j$*.*

The classical Sklar's Theorem (Sklar 1959) highlights how multivariate copulas introduced in Definition 2.5 model the dependence structure between random variables. For the sake of completeness, we report here the enunciation of the theorem, adapted to our case:

Theorem 2.6 (sklar1959)**.** *Let* $F_{i_1,\dots,i_s}$ *be the joint distribution function of the s-ple* $(t_{i_1}, \dots, t_{i_s})$*, with* $i_1, \dots, i_s = 1, \dots, m$*. Define the margins as* $F_{i_1}, \dots, F_{i_s}$*. Then there exists a s-variate copula* $C_{i_1,\dots,i_s}$ *such that, for each* $x_1, \dots, x_s \in \mathbb{R}$*,*

$$F_{i_1,\dots,i_s}(x_1, \dots, x_s) = C_{i_1,\dots,i_s}(F_{i_1}(x_1), \dots, F_{i_s}(x_s)). \tag{1}$$

If the margins $F_{i_1}, \dots, F_{i_s}$ *are continuous, then the copula* $C_{i_1,\dots,i_s}$ *is unique. Conversely, if* $C_{i_1,\dots,i_s}$ *is a s-variate copula and* $F_{i_1}, \dots, F_{i_s}$ *are distribution functions, then the function* $F_{i_1,\dots,i_s}$ *defined in (1) is a s-dimensional distribution function with margins* $F_{i_1}, \dots, F_{i_s}$*.*

Theorem 2.6 points out that, given a set of random variables, the relationship between joint and marginal distributions is stated through copulas.

In the presence of bivariate copulas, i.e. for $s = 2$, we can derive the covariance between couples of random variables:

Proposition 2.7.

$$\sigma_{i,j} = \frac{1}{\sigma_i \sigma_j} \int\int_{\mathbb{R}^2} \left[C_{i,j}(F_i(x), F_j(y)) - F_i(x)F_j(y)\right] \mathrm{d}x\mathrm{d}y,$$

where $C_{i,j}$ is the bivariate copula defined as in (1).

Proposition 2.7 shows that the covariance between couples of random variables (t_i, t_j) can be derived from the knowledge of the copula describing their stochastic dependence. More generally, we can say that the introduction of a multivariate copula leads to the identification of a variance-covariance matrix.

3. Main results

The argument on the stochastic dependence developed above can be applied to the FDR control of the multiple statistics $\boldsymbol{t}$. In order to proceed, we need a condition on the sets $\boldsymbol{p}^s$ introduced above:

Condition 3.1. *One of the following assumptions holds:*
(A3.1.i) if $p_i \in \boldsymbol{p}_0$, then there exists a unique $s_i \in \{1, \ldots, S\}$ such that $p_i \in \boldsymbol{p}^{s_i}$. Moreover, for each $s = 1, \ldots, S$, it must be:

$$m^s = \begin{cases} 2, & \text{if } \boldsymbol{p}_0^s \neq \emptyset; \\ \text{arbitrary}, & \text{otherwise.} \end{cases}$$

(A3.1.ii) $\boldsymbol{p}^{s_i} \cap \boldsymbol{p}^{s_j} = \emptyset$, for $s_i \neq s_j$, and $m^s = 2$, for each $s = 1, \ldots, S$.

Condition 3.1 means that the division of the set $\boldsymbol{p}$ in the subsets $\boldsymbol{p}^s$ is such that each p value of a true null hypothesis is contained in one $\boldsymbol{p}^s$, and each $\boldsymbol{p}^s$ containing a p value of a true null hypothesis has cardinality equals to 2. This is not a restrictive hypothesis, since the decomposition of $\{\boldsymbol{p}^s\}_{s=1,\ldots,S}$ to be used for the ssBH procedure can be arbitrarily chosen. It is worth noting that when (A3.1.ii) holds, then $m^s = 2$, for each $s = 1, \ldots, S$; if (A3.1.i) is true, then $\exists\, \tilde{S} \leqslant S$ such that $m^s = 2$, for each $s = 1, \ldots, \tilde{S}$.

We are now able to state our first main result:

Proposition 3.2. *Assume that Condition 3.1 holds and that the dependence between the statistics in $\boldsymbol{t}^s$ is described by a copula C_s such that:*

$$C_s(u, v) = uv + \theta\phi(u)\phi(v), \tag{2}$$

for each $s = 1, \ldots, S$, with $\theta \in [-1, 1]$ and $\phi : [0, 1] \to [0, 1]$ satisfying the following conditions:
(C3.2.i) $\phi(0) = \phi(1) = 0$;
(C3.2.ii) ϕ is Lipschitzian in $[0, 1]$, i.e.: $|\phi(x) - \phi(y)| \leqslant |x - y|$, for each $x, y \in [0, 1]$;
(C3.2.iii) ϕ is convex or concave in $[0, 1]$.
Then the level q ssBH procedure controls the FDR at level qm_0/m.

The main limitations of this approach are basically two. First, Proposition 3.2 refers to *couples*, i.e. subsets of cardinality 2; furthermore, copula in (2) is *symmetric* with respect to its arguments, i.e.: $C_s(u,v) = C_s(v,u)$ for each u, v. The latter aspect is a very strong requirement for multivariate modelling, in that symmetric copulas are able to cover only a small range of dependencies.

However, it is possible to generalize the result by using a s-variate approach, with $s > 2$, in a not necessarily symmetric framework. A copula approach will be adopted also in this case to ensure FDR control.

We first extend the analysis to cover the multivariate case, with $s > 2$. Then we provide a generalization to the asymmetric setting.

To deal with the s-variate symmetric framework, it is useful to recall two definitions:

Definition 3.3. *Consider a continuous strictly decreasing convex function*

$$\psi : [0,1] \to [0,+\infty) \tag{3}$$

such that $\psi(1) = 0$ *and* $\lim_{x\to 0^+} \psi(x) = +\infty$.

An s-variate Archimedean copula with generator ψ is a copula $C_s^{(\psi)}$ such that

$$C_s^{\psi}(u_1,\ldots,u_s) = \psi^{-1}\left(\sum_{i=1}^{s} \psi(u_i)\right). \tag{4}$$

Analogously to what already noted for copula in (2), also copula C_s^{ψ} in (4) refers to a symmetric case.

We now introduce a generalization of the monotonic property for functions:

Definition 3.4. *A function*

$$\psi : [0,1] \to [0,+\infty) \tag{5}$$

is completely monotone in $[0,1]$ *if and only if* $\psi \in C^{\infty}(0,1) \cap C^{0}[0,1]$, *and* $(-1)^n \psi^{(n)}(x) \geqslant 0, \forall\, n = 0,1,2,\ldots;\ \forall\, x \in (0,1)$.

The following result states a sufficient condition for FDR control in the s-variate symmetric case:

Proposition 3.5. *Assume that the dependence between the statistics in $\boldsymbol{t}^s$ is described by an Archimedean copula C_s^{ψ}, i.e.:*

$$\begin{cases} C_s^{\psi}(u_1,\ldots,u_s) = \psi^{-1}\left(\sum_{i=1}^{s} \psi(u_i)\right); \\ u_k = F_k(x_k), \qquad x_k \in \mathbb{R},\ \forall\, k = 1,\ldots,s. \end{cases}, \tag{6}$$

where ψ is completely monotone in [0,1].

Then the level q BH procedure controls the FDR at level qm_0/m.

The generalization to the asymmetric framework can be obtained at the cost of some mildly stronger assumptions. We enter the details, by firstly introducing an asymmetric copula constituting a generalization of the Archimedean copula proposed in (4).

Definition 3.6. *Let us introduce a set of $s \times m$ functions*

$$h_{jk} : [0,1] \to [0,1], \qquad j = 1,\dots,m; \;\; k = 1,\dots,s \tag{7}$$

such that:
(C3.6.i) h_{jk} is differentiable and strictly increasing, for each j,k;
(C3.6.ii) $h_{jk}(0) = 0$ and $h_{jk}(1) = 1$;
(C3.6.iii) $\frac{1}{m}\sum_{j=1}^{m} h_{jk}(x) = x$, for each $k = 1,\dots,s$ and $x \in [0,1]$.
Moreover, define

$$\psi : [0,1] \to [0,1] \tag{8}$$

such that:
(C3.6.iv) ψ is $s+2$ times differentiable in $(0,1)$;
(C3.6.v) $\psi^{(i)} > 0$, for $i = 1,\dots,s$;
(C3.6.vi) $\psi(0) = 0$ and $\psi(1) = 1$. We define an (Archimedean) asymmetric copula as C^{ψ}_{AS} : $[0,1]^s \to [0,1]$ such that:

$$C^{\psi}_{AS}(u_1,\dots,u_s) = \psi^{-1}\left(\frac{1}{m}\sum_{j=1}^{m}\prod_{k=1}^{s} h_{jk}(\psi(u_k))\right). \tag{9}$$

Copula in (9) has been firstly introduced and explored in Liebscher (2008). It is worth noting that, as far as the copula's definition is concerned, conditions (C3.6.iv) and (C3.6.v) could be weakened, our stronger version being required to prove the following general result:

Proposition 3.7. *Assume that*

$$\left(\psi^{-1}\right)^{(s+2)}(x)\left(\psi^{-1}\right)^{(s)}(x) - \left[\left(\psi^{-1}\right)^{(s+1)}(x)\right]^2 \geqslant 0, \qquad \forall\, x \in (0,1) \tag{10}$$

and

$$\psi(u_k) = h^{-1}_{jk}\left(\frac{e^{u_k}-1}{e-1}\right), \qquad j = 1,\dots,m; \;\; k = 1,\dots,s. \tag{11}$$

Moreover, suppose that the dependence between the statistics in $\boldsymbol{t}^s$ is described by copula (9). Then the level q BH procedure controls the FDR at level q_0/m.

4. Discussion

The introduction of copulas to model stochastic dependence allows us to deal with more general dependence structures than Pearson's correlation or those based on linearity, which are appropriate only when referring to normal multivariate models. In this respect, it is worth noting that covariance can be derived directly from copulas (see Proposition 2.7), but the converse is not true.

In our framework, Propositions 3.2, 3.5, and 3.7 provide *sufficient* conditions for the FDR to hold in the presence of fairly general dependence schemes.

Proposition 3.2 offers viable ways of selecting the couples in such a way that the conditions for validly using the ssBH procedure are satisfied. The main limitations are related to the cardinality of the subsets and to the condition of symmetry between couples of random variables. As a good feature, we must notice that the pairwise dependence introduced in the set up given by family of $\{\boldsymbol{p}^s\}_{s=1,\dots,S}$, Condition 3.1 and the copulas in (2) allow us to describe a system with both positively and negatively correlated test statistics. Indeed, the positive dependence condition formalized by TP2 (see Proposition 2.4) is required only for some pairs of statistics in $\boldsymbol{t}$ (the ones appearing in the $\boldsymbol{t}^s$'s), while no assumptions are stated on the remaining couples. This aspect meets a natural requirement on the dependence structure of statistics in multiple testing.

Furthermore, copula defined in (2) allows to derive an explicit expression for the correlation between the individual statistics in $\boldsymbol{t}^s$. Indeed, some algebra provides that if the stochastic dependence between X and Y is described through copula C_s in (2), then the correlation coefficient $\rho_{X,Y}$ between X and Y can be written as:

$$\rho_{X,Y} = 12\theta \left(\int_0^1 \phi(\xi) d\xi \right)^2 . \tag{12}$$

Notice also that copula (2) used in Proposition 3.2 is a "perturbation" of the product copula: when $\theta = 0$ the case collapses to independence. It is also worth noting that copula in (2) is a generalization of the Farlie-Gumbel-Morgestern (FGM) copula that holds when $\phi(u) = u(1-u)$. In the bivariate case, when $\theta \geqslant 0$, it implies *positive quadrant dependence* (see Lai and Xie 2000), which is a weaker form of dependence than TP2. However, a word of caution is in order here. The FGM copula, as well as its studied variants, are known for implying only modest dependence (see, e.g., Huang and Kotz 1999): therefore, we cannot expect the copula (2) to accurately represent very strong dependence across the test statistics. As far as the "pure" FGM copula is concerned, its dependence as measured by Kendall's τ and Spearman's ρ is respectively $2\theta/9$ and $\theta/3$ with $-1 \leqslant \theta \leqslant 1$. However, an appropriate choice of a function $\phi(u) \neq u(1-u)$ satisfying conditions in Proposition 3.2 allows to strengthen the typical weak dependence structure of the FGM copula.

The need to obtain a truly multivariate result motivates the formulation of Proposition 3.5. As with the bivariate case, it is worth noting that the introduction of the family of sets $\{\boldsymbol{p}^s\}_{s=1,\dots,S}$ with the stochastic dependence structure formalized in Proposition 3.5 allows us to describe a system with both positively and negatively correlated test statistics. Indeed, the positive dependence condition formalized by MTP2 is required only for some statistics in $\boldsymbol{t}$ (the ones appearing in the $\boldsymbol{t}^s$'s), while no assumptions are stated on the remaining statistics. Furthermore, the Archimedean copulas (6) used in Proposition 3.5 are more flexible than the particular case of FGM copulas in that they can represent cases with both strong positive and negative dependence. Kendall's τ for Archimedean copulas takes the convenient form (see Genest and MacKay 1986, Theorem 2)

$$\tau_{X,Y} = 4 \int_0^1 \frac{\psi(\xi)}{\psi'(\xi)} d\xi + 1 . \tag{13}$$

Copulas in (2) and in (6) exhibit a symmetry property, in that they are invariant with respect to permutation of their univariate arguments. Such a symmetry is able to model dependence structures which depend only on a small number of parameters. This is their main limitation, in that they are not particularly flexible in fitting multivariate data with a large number of parameters. It is also worth noting that symmetric copulas are able to model only a rather small range of dependencies.

For these reasons, the extension to the asymmetric case has been proposed. Proposition 3.7 extends the FDR applicability to situations where dependence can be well represented by asymmetric copulas (see Liebscher 2008, 2011, on asymmetric copulas). In this respect, Proposition 3.7 complements and extends Yekutieli (2008).

5. Proofs

In this section we provide the proofs for our main results.

Proposition 3.2

Proof. Conditions (C3.2.i) and (C3.2.ii) guarantee that C_s in (2) is a copula.

Denote as X and Y the individual statistics in $\boldsymbol{t}^s$. Amblard and Girard (2002) shows that, if the dependence between X and Y is described through the copula C_s in (2) and condition (C3.2.iii) holds, then Y is stochastically increasing in X and X is stochastically increasing in Y, i.e. the following conditions hold:

$$\begin{cases} P(Y > y \,|\, X = x) \text{ is nondecreasing in } x, \ \forall\, y; \\ P(X > x \,|\, Y = y) \text{ is nondecreasing in } y, \ \forall\, x. \end{cases} \tag{14}$$

The system (14) is equivalent to the TP2 property for the set $\boldsymbol{t}^s$ (see Nelsen 2006). Hence, Condition 3.1 and Proposition 2.2 in Yekutieli (2008) give the thesis. □

Proposition 3.5

Proof. Denote as $t_1, \ldots, t_s$ the individual statistics in $\boldsymbol{t}^s$. If the stochastic dependence in $\boldsymbol{t}^s$ is described as in the hypotheses, then Müller and Scarsini (2005) guarantees that the MTP2 property holds for $t_1, \ldots, _s$. Hence, Proposition 2.2 in Yekutieli (2008) gives the thesis. □

Proposition 3.7

Proof. We need to prove that C^{ψ}_{AS} satisfies the MTP2 property. Then, Proposition 2.2 in Yekutieli (2008) gives the thesis.

In virtue of Müller and Scarsini (2005), it is sufficient to check that the density f of C^{ψ}_{AS} is log-supermodular, that is equivalent to say that

$$\log(f(u_1, \ldots, u_s)) := \log\left(\frac{\partial^s}{\partial u_1 \ldots \partial u_s} C^{\psi}_{AS}(u_1, \ldots, u_s)\right) \tag{15}$$

is supermodular.

By (9) we have

$$\begin{aligned} f(u_1,\dots,u_s) &= \frac{\partial^s}{\partial u_1\dots\partial u_s} C_{AS}^{\psi}(u_1,\dots,u_s) \\ &= \left(\psi^{[-1]}\right)^{(s)} \left(\frac{1}{m}\sum_{j=1}^{m}\prod_{k=1}^{s} h_{jk}(\psi(u_k))\right) \times \\ &\quad \times \frac{1}{m}\sum_{j=1}^{m}\prod_{k=1}^{s} h'_{jk}(\psi(u_k))\psi'(u_k)\,. \end{aligned} \tag{16}$$

By (16) we can write

$$\begin{aligned} \log(f(u_1,\dots,u_s)) &= \log\left[\left(\psi^{[-1]}\right)^{(s)}\left(\frac{1}{m}\sum_{j=1}^{m}\prod_{k=1}^{s} h_{jk}(\psi(u_k))\right)\right] + \\ &\quad + \log\left[\frac{1}{m}\sum_{j=1}^{m}\prod_{k=1}^{s} h'_{jk}(\psi(u_k))\psi'(u_k)\right] \\ &=: A(u_1,\dots,u_s) + B(u_1,\dots,u_s)\,, \end{aligned} \tag{17}$$

where the terms $A(\cdot)$ and $B(\cdot)$ are an intuitive shorthand for the two $\log[\cdot]$ terms.

The supermodularity of $\log[f(u_1,\dots,u_s)]$ is equivalent to the following condition:

$$\frac{\partial^2}{\partial u_{k_1}\partial u_{k_2}}\left[A(u_1,\dots,u_s) + B(u_1,\dots,u_s)\right] \geqslant 0, \tag{18}$$

for each $k_1, k_2 \in \{1,\dots,s\}$, and $(u_1,\dots,u_s) \in [0,1]^s$. For an easier notation, we will pose hereafter

$$\xi := \frac{1}{m}\sum_{j=1}^{m}\prod_{k=1}^{s} h_{jk}(\psi(u_k)). \tag{19}$$

We analyse the terms $A(\cdot)$ and $B(\cdot)$ separately.

First notice that

$$\frac{\partial A(u_1,\dots,u_s)}{\partial u_{k_1}} = \frac{\left(\psi^{[-1]}\right)^{(s+1)}(\xi)}{\left(\psi^{[-1]}\right)^{(s)}(\xi)} \times \frac{1}{m}\sum_{j=1}^{m} h'_{jk_1}(\psi(u_{k_1}))\psi'(u_{k_1}) \prod_{k\neq k_1}\left[h_{jk}(\psi(u_k))\right]$$

and

$$\begin{aligned}\frac{\partial^2 A(u_1,\dots,u_s)}{\partial u_{k_1}\partial u_{k_2}} &= \left\{\frac{1}{m}\sum_{j=1}^{m} h'_{jk_1}(\psi(u_{k_1}))\psi'(u_{k_1})\prod_{k\neq k_1}[h_{jk}(\psi(u_k))]\right\}\times\\ &\times\left\{\frac{1}{m}\sum_{j=1}^{m} h'_{jk_1}(\psi(u_{k_2}))\psi'(u_{k_2})\prod_{k\neq k_2}[h_{jk}(\psi(u_k))]\right\}\times\\ &\times\left\{\left(\psi^{[-1]}\right)^{(s+2)}(\xi)\times\left(\psi^{[-1]}\right)^{(s)}(\xi)-\left[\left(\psi^{[-1]}\right)^{(s+1)}(\xi)\right]^2\right\}+\\ &+\left(\psi^{[-1]}\right)^{(s)}(\xi)\times\left(\psi^{[-1]}\right)^{(s+1)}(\xi)\times\\ &\times\frac{1}{m}\sum_{j=1}^{m} h'_{jk_1}(\psi(u_{k_1}))\psi'(u_{k_1})h'_{jk_2}(\psi(u_{k_2}))\psi'(u_{k_2})\times\\ &\times\prod_{k\neq k_1,k_2}[h_{jk}(\psi(u_k))]\ .\end{aligned}\tag{20}$$

Hence, under condition (C3.6.v) and hypothesis (10), we have

$$\frac{\partial^2 A(u_1,\dots,u_s)}{\partial u_{k_1}\partial u_{k_2}}\geqslant 0.\tag{21}$$

Let us now turn to $B(\cdot)$:

$$\frac{\partial B(u_1,\dots,u_s)}{\partial u_{k_1}}=\frac{\sum_{j=1}^m\left[h''_{jk_1}(\psi(u_{k_1}))\psi'(u_{k_1})+h'_{jk_1}(\psi(u_{k_1}))\psi''(u_{k_1})\right]\prod_{k\neq k_1}h'_{jk}(\psi(u_k))\psi'(u_k)}{\sum_{j=1}^m\prod_{k=1}^s h'_{jk}(\psi(u_k))\psi'(u_k)},$$

hence we have:

$$\begin{aligned}\frac{\partial^2 B(u_1,\dots,u_s)}{\partial u_{k_1}\partial u_{k_2}} &= -\frac{\sum_{j=1}^m\prod_{k\neq k_1,k_2}h'_{jk}(\psi(u_k))\psi'(u_k)}{\left[\sum_{j=1}^m\prod_{k=1}^s h'_{jk}(\psi(u_k))\psi'(u_k)\right]^2}\times\\ &\times\Big\{\prod_{k=k_1,k_2}[h''_{jk}(\psi(u_k))(\psi'(u_k))^2+h'_{jk}(\psi(u_k))\psi''(u_k)]\times\\ &\times\frac{1}{m}\sum_{j=1}^m\prod_{k=1}^s h'_{jk}(\psi(u_k))\psi'(u_k)-\\ &-\left[h''_{jk_1}(\psi(u_{k_1}))(\psi'(u_{k_1}))^2+h'_{jk_1}(\psi(u_{k_1}))\psi''(u_{k_1})\right]\times\\ &\times h'_{jk_2}(\psi(u_{k_2}))\psi'(u_{k_2})\times\\ &\times\frac{1}{m}\sum_{j=1}^m\left[h''_{jk_2}(\psi(u_{k_2}))(\psi'(u_{k_2}))^2+h'_{jk_2}(\psi(u_{k_2}))\psi''(u_{k_2})\right]\times\\ &\times\prod_{k\neq k_2}h'_{jk}(\psi(u_k))\psi'(u_k)\Big\}.\end{aligned}\tag{22}$$

By (22) we obtain that a sufficient condition for being $\partial^2 B(u_1,\dots,u_s)/(\partial u_{k_1}\partial u_{k_2})=0$ is:

$$h''_{jk_2}(\psi(u_{k_2}))(\psi'(u_{k_2}))^2+h'_{jk_2}(\psi(u_{k_2}))\psi''(u_{k_2})=h'_{jk_2}(\psi(u_{k_2}))\psi'(u_{k_2}).\tag{23}$$

We then need to solve the ODE in (23), with the initial condition given by (C3.6.vi).
Note that (23) is equivalent to

$$\frac{h''_{jk_2}(\psi(u_{k_2}))\psi'(u_{k_2})}{h'_{jk_2}(\psi(u_{k_2}))} = -\frac{\psi''(u_{k_2}) - \psi'(u_{k_2})}{\psi'(u_{k_2})},$$

which leads to:

$$\log[h'_{jk_2}(\psi(u_{k_2}))] = -\int \frac{\psi''(u_{k_2}) - \psi'(u_{k_2})}{\psi'(u_{k_2})} \mathrm{d}u_{k_2}.$$

Then

$$h'_{jk_2}(\psi(u_{k_2})) = \frac{H_1 \mathrm{e}^{u_{k_2}}}{\psi'(u_{k_2})}, \qquad H_1 \in (0, +\infty). \tag{24}$$

(24) can be rewritten as follows:

$$h'_{jk_2}(\psi(u_{k_2})) \cdot \psi'(u_{k_2}) = H_1 \mathrm{e}^{u_{k_2}},$$

which immediately leads to

$$h_{jk_2}(\psi(u_{k_2})) = H_1 \mathrm{e}^{u_{k_2}} + H_2, \qquad H_2 \in \mathbb{R}.$$

Since h_{jk_2} is invertible, by imposing the initial conditions in (C3.6.vi) we obtain:

$$\psi(u_{k_2}) = h^{-1}_{jk_2}\left(\frac{\mathrm{e}^{u_{k_2}} - 1}{\mathrm{e} - 1}\right),$$

that is exactly condition (11) when $k = k_2$.
The result is proved, by the arbitrariness of k_2. □

Acknowledgements

We thank Rachele Foschi and Fabio Spizzichino for comments and discussion. Materials related to this paper were presented at the Conference in honour of Professor M.H. Pesaran (Cambridge, July 2011): we are grateful to conference participants for their constructive comments.

References

Amblard C, Girard S (2002). "Symmetry and Dependence Properties within a Semiparametric Family of Bivariate Copulas." *Nonparametric Statistics*, **14**(6), 715–727.

Benjamini Y, Hochberg Y (1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society, Series B*, **57**(1), 289–300.

Benjamini Y, Yekutieli D (2001). "The Control of the False Discovery Rate in Multiple Testing under Dependency." *Annals of Statistics*, **29**(4), 1165–1188.

Genest C, MacKay J (1986). "The Joy of Copulas: Bivariate Distributions with Uniform Marginals." *The American Statistician*, **40**(4), 280–283.

Huang JS, Kotz S (1999). "Modifications of the Farlie-Gumbel-Morgenstern Distributions. A Tough Hill to Climb." *Metrika*, **49**(2), 135–145.

Lai CD, Xie M (2000). "A New Family of Positive Quadrant Dependent Bivariate Distributions." *Statistics & Probability Letters*, **46**(4), 359–364.

Liebscher E (2008). "Construction of Asymmetric Multivariate Copulas." *Journal of Multivariate Analysis*, **99**(10), 2234–2250.

Liebscher E (2011). "Erratum to "Construction of asymmetric multivariate copulas" [J. Multivariate Anal. 99 (2008) 2234–2250]." *Journal of Multivariate Analysis*, **102**(4), 869–870.

Müller A, Scarsini M (2005). "Archimedean Copulae and Positive Dependence." *Journal of Multivariate Analysis*, **93**(2), 434–445.

Nelsen RB (2006). *An Introduction to Copulas*. Springer Series in Statistics, 2nd edition. Springer, New York.

Sklar A (1959). "Fonctions de répartition à n dimensions et leurs marges." *Publications de l'Institut de Statistique de L'Université de Paris*, **8**, 229–231.

Sorić B (1989). "Statistical 'Discoveries' and Effect-Size Estimation." *Journal of the American Statistical Association*, **84**(406), 608–610.

Yekutieli D (2008). "False Discovery Rate Control for Non-positively Regression Dependent Test Statistics." *Journal of Statistical Planning and Inference*, **138**(2), 405–415.

Affiliation:

Roy Cerqueti
University of Macerata, Dept. of Economic and Financial Institutions.
Via Crescimbeni, 20. I-62100 Macerata, Italy E-mail: roy.cerqueti@unimc.it

Mauro Costantini
Brunel University, Dept. of Economics and Finance.
Kingston Lane, Uxbridge. Middlesex UB8 3PH, United Kingdom
E-mail: Mauro.Costantini@brunel.ac.uk

Claudio Lupi
University of Molise, Dept. of Economics, Management, and Social Sciences.
Via De Sanctis. I-86100 Campobasso, Italy
E-mail: lupi@unimol.it